

The Performance of UK Investment Trusts

YASEMIN BAL and LAWRENCE A. LEGER

The Performance of 92 UK investment trusts was analysed over the period 1975 to 1993 using the Sharpe Treynor and Jensen measures of portfolio performance. A very high degree of correlation was found between the measures. Even without correction for transactions costs funds did not on average outperform the market, although a few individual funds appeared to do so. Fund rankings by the Sharpe measure showed significant intertemporal persistence, especially in the income-producing group of funds, which needs further investigation.

INTRODUCTION

The ability of fund managers to select winning stocks, otherwise referred to as 'selectivity', has been the subject of much research since the seminal work of Treynor [1965], Sharpe [1966] and Jensen [1968], who developed methods for risk-adjustment in fund evaluation based on mean-variance analysis. Despite ambiguities in the interpretation of these measures they continue to be widely used. For example, recent studies on US data using these measures include Ippolito [1989], Cumby and Glen [1990], Eun, Kolodny and Resnik [1991] and Cheney Atkinson and Bailey [1992]. A typical conclusion of studies based on these traditional performance measures is that investment companies have not been able to outperform the market consistently on a risk-adjusted basis net of transactions costs.

A few researchers [for example Lakonishok *et al.*, 1992; Grinblatt and Titman, 1992] have studied the persistence of fund performance over various time intervals, with mixed results. Evidence of persistent above-average performance by particular funds would be not only inconsistent with efficient markets, but of considerable practical value to investors. Indeed, a great deal of advertising¹ by managed funds emphasises their claims to persistent high rankings. In a large-scale study of the US fund management industry Lakonishok *et al.* [1992] suggest that there may some

Yasmin Bal is at the Turkish Prime Ministry Under-Secretariat for Treasury and Foreign Trade and the University of Bilkent, Ankara, Turkey.

Lawrence Leger is at Loughborough University of Technology, Department of Economics, Loughborough University, LE11 3TU, UK.

weak persistence in fund performance over two and three-year periods (although returns did not in general exceed the S & P 500 after transactions costs). Grinblatt and Titman [1992] also found some evidence of persistent abnormal performance for US mutual funds, using an 8-portfolio benchmark developed earlier [Grinblatt and Titman, 1989] as an extension to Jensen [1968].

There has been little evaluation of UK funds and even less which explicitly compares the Sharpe, Treynor and Jensen methodologies. The study by Cheney, Atkinson and Bailey [1992] compared UK international trusts with US and Canadian funds, but this study was limited to 20 trusts and a two-year sample period 1988–90.

We have attempted to redress the lack of research on UK funds by comparing the traditional methods of Treynor [1965], Sharpe [1966] and Jensen [1968]. We use a much larger sample than that of Cheney *et al.* [1992], including 92 UK investment trusts evaluated over 14 five-year rolling periods from 1975 to 1993. We also examine both the correlation between different measures of performance and the persistence in fund performance over time. With regard to fund performance, we find that although the Jensen measure of excess return averaged across funds was positive for ten of the fourteen periods, the measure fails to reach significance for the great majority of individual funds. With regard to comparative evaluation, we find a high degree of correlation between performance measures, with some degree of significant inter-temporal consistency in the ranking of the funds by the Sharpe measure.

THE MODELS

All three measures are based on mean-variance analysis, but Sharpe [1966] measured the total risk of funds while both Treynor [1965] and Jensen [1968] adjusted for systematic (beta) risk. Sharpe's measure of selectivity, S_p , is given by

$$(1) \quad S_p = \frac{\bar{R}_p - \bar{R}_f}{\sigma_p}$$

where \bar{R}_p is mean return to portfolio p during the evaluation period, \bar{R}_f is the mean risk free rate of return and σ_p is the standard deviation of the portfolio return. S_p is therefore the slope of a straight line connecting \bar{R}_f with the point (σ_p, \bar{R}_p) . A higher value of S_p indicates a higher risk-adjusted return, allowing a ranking of funds.

S_p does not of itself measure an excess return, since this can be assessed only with respect to some benchmark portfolio (defined independently), but

it does have another useful interpretation. In the framework of the Capital Asset Pricing Model [CAPM] and two-fund separation an efficient portfolio is a linear combination of the risk-free rate and the market portfolio and lies on the Capital Market Line (where S_p is maximised). Hence S_p is a measure of how efficiently a fund is diversified. Note, however, that if a fund manager is good at selecting winners but the fund is not well diversified, *ex-post* higher returns from good asset selection can offset lower risk-adjusted return arising from the poor diversification.

Treynor [1965] uses systematic risk, because the portfolio under evaluation is assumed to be part of a larger portfolio in which non-systematic risk is diversified away. (The Sharpe and Treynor measures will give consistent rankings if funds are well-diversified.) The Treynor measure is given by

$$(2) \quad S_p = \frac{\bar{R}_p - \bar{R}_f}{\beta_p}$$

where \bar{R}_p is the mean portfolio return, \bar{R}_f is the mean risk free rate of return and β_p is systematic risk, estimated by the market model of equation (3):

$$(3) \quad \bar{R}_p = \alpha_p + \beta_p \bar{R}_m + \epsilon_p.$$

\bar{R}_m is the return to the market and ϵ_p is a white noise error term.

Rather than ranking funds, Jensen [1968] suggests that the statistical significance of excess returns can be estimated from an *ex-post* form of the CAPM based on the market model:

$$(4) \quad \bar{R}_{pt} - \bar{R}_{ft} = \alpha_p + \beta_p (\bar{R}_{mt} - \bar{R}_{ft}) + u_{pt} \quad \forall p.$$

Here α_p is Jensen's differential performance measure for fund p , showing the manager's ability to select winners, and u_{pt} is a white noise term. Jensen calculates the statistical significance of deviations of an estimated α_p from its CAPM-predicted value of zero.

INTERPRETING THE MEASURES

There are a number of difficulties in the interpretation of these traditional measures. Some of the more important issues are briefly discussed below.

There is an implicit benchmark portfolio in Treynor's analysis (unlike that of Sharpe) because T_p can only be estimated with respect to a market index, making it difficult to interpret the measure within a CAPM equilibrium framework. If the chosen market index were efficient then all

funds would lie on the implied security market line [following Roll, 1977], so that T_p would necessarily be the same for all funds in equilibrium, vitiating attempts at performance evaluation. If, on the other hand, T_p were found to discriminate between funds, it would imply that the proxy to the market portfolio was mean-variance inefficient, and possibly a rejection of the CAPM, thereby raising doubts about the suitability of the implied risk-adjustment. Furthermore, since different benchmarks could produce different estimates of beta it is theoretically possible for Treynor rankings to be altered by different choices of index [see also Ross, 1978]. Fortunately, tests of the CAPM suggest [see Stambaugh, 1982] that results are generally insensitive to the choice of index, while Dybvig and Ross [1985] have shown that funds which plot above or below a security market line generated by an inefficient index will still do so when the index is changed, as long as a risk-free asset exists.

The problems raised by inefficient benchmarks also apply to the Jensen measure. If the CAPM is true then no fund should earn an excess return when this is estimated over an appropriate sample period [α_p should not deviate from its theoretical value] but if the CAPM is untrue there is difficulty in accepting the implied benchmark [see Roll, 1977 and 1978]. Admati and Ross [1985] note that one resolution of the problem, implied by the traditional models, is to maintain the homogeneous expectations assumption of the CAPM and accept that fund managers may trade on private information without influencing the homogeneous expectations equilibrium. They object to this view and show that if equilibrium is based on heterogeneous expectations with informational asymmetries between and among market observers and fund managers then the use of traditional single index models based on homogeneous expectations may be inappropriate and misleading. A more pragmatic approach, followed here, is to acknowledge the chosen market index as an imperfect benchmark which is nonetheless informative and useful in careful practice.

It has also been shown [Grant, 1977] that the Jensen measure will be biased downward if managers attempt to 'time' their investments (switching between low- and high-beta stocks according to market conditions). The substantive empirical significance of this bias seems to be still unresolved – Chang and Llewellyn [1984] were unable to demonstrate an increased frequency (an increase only from 4 to 5 funds, out of a total of 67) of statistically significant Jensen measures when comparing the traditional model with a more sophisticated 'timing' methodology; Lee and Rahman [1990] found a bigger increase (an increase from 13 to 24, out of a total of 93), but with excess performance more likely to be negative than positive. Other studies on market timing include Treynor and Mazuy [1967], Kon and Jen [1979], Kon [1983], Henriksson [1894], Grinblatt and Titman [1989],

Coggin *et al.* [1993].

The Treynor and Sharpe measures both require careful interpretation when the empirical market lines are negatively-sloped. In this context we find that when two funds have same mean return the fund with the greater risk will have the greater Sharpe or Treynor measure. It therefore makes sense to examine the absolute values of the measures. It is also possible for the Sharpe and Treynor measures to give different fund rankings – of two funds with the same mean return the fund with the lower systematic risk could have the higher total risk.

Finally we note that some researchers have attempted to overcome the problem of benchmark inefficiency by moving away from the single-index framework altogether. Examples of studies using multi-index/Arbitrage Pricing Model benchmarks include Connor and Korajczyk [1986], Lehman and Modest [1987], and Grinblatt and Titman [1992]. Here we note that such models bring their own problems, since the various risk factors are not easy either to identify or interpret, unlike the single market index of the traditional approach.

PERFORMANCE OF FUNDS

We compare each of the above models while evaluating the performance of 92 UK investment trusts and the Financial Times All-Share Index over 14 five-year rolling periods from 1975 to 1993. Monthly returns were calculated from share prices adjusted² for capital issues and large once-off payments. The performance of funds as a whole was assessed by comparing equally-weighted portfolio averages of all 92 funds with the market index for each of the sample periods, for both risk-adjusted and risk-unadjusted performance.³ The statistical significance of the Jensen selectivity measure was calculated for each fund, α_p being estimated by OLS, with a Cochran-Orcutt GLS procedure used to correct for any residual autocorrelation detected by the Durbin Watson Statistic. The degree of agreement between the Sharpe and Treynor measures was estimated by Spearman's rank-order correlation coefficient, ρ [see Kendall, 1970]. Persistence in the Sharpe ranking across non-overlapping periods was also assessed using this test.

(i) Overall Fund Performance compared to the Market Index

Summary statistics⁴ for overall fund performance during each rolling period are presented in Table 1, showing measures averaged across all funds compared to the same measures on the market index. A non-risk-adjusted comparison shows that funds as a whole outperformed the market index in 11 of the 14 rolling periods, with the market showing superior mean return in first period and in the last three periods (where the empirical capital

market line is negatively sloped).

The Sharpe and Treynor risk-adjusted performance measures are contradictory. Using the absolute values of the measures as our criteria we find that under Sharpe the market index outperformed the funds for 8 of the 14 periods, but under Treynor the funds outperformed the market in all periods. Since the Sharpe and Treynor measures do not allow the statistical significance to be assessed, our evaluation of abnormal performance is deferred to the next section. We note, however, that funds were not in general well-diversified, so that the Treynor index could be the more appropriate ranking device for investors able to diversify for themselves. To examine diversification, note that risk is partitioned into systematic and non-systematic risk under the single-index restriction⁵ on the market model, giving

$$(5) \quad \sigma_p^2 = \beta_p^2 \sigma_m^2 + \sigma_{\varepsilon p}^2$$

so that $S_p \sigma_m / T_p$ is necessarily unity for funds which are well-diversified (implying zero non-systematic risk) with respect to the index. Average fund diversification scores for the 14 periods are therefore estimated by $S_p \sigma_m / T_p$ from the data in Table 1 and are given in Table 2. Values less than unity⁶ are evident in all cases.

(ii) Statistical Significance of Excess Returns

Neither S_p nor T_p assesses the statistical significance of fund performance – both are primarily devices for ranking portfolios. To assess significance we rely on Jensen's α_p . Table 1 reveals that the simple average of the Jensen measure across funds is positive in 11 of the 14 periods, agreeing with the non-risk-adjusted results. However, when α_p for individual funds is examined it is apparent that very few funds achieve significant excess returns in any period.

TABLE 1
SHARPE, TREYNOR AND JENSEN MEASURES: FUND AVERAGE RETURN VERSUS THE
MARKET INDEX AND A RISK-FREE PROXY, 1975-93

Period	Averages over all Funds				Market			Riskless Asset
	Mean HPR [%]	Sharpe Measure	Treynor Measure	Jensen Measure	Mean HPR [%]	Sharpe Measure	Treynor Measure	Mean HPR [%]
75-80	2.3151	0.1289	2.1826	-.0470	2.4654	0.1745	1.67	.7954
76-81	1.3946	0.0608	0.7918	.23015	1.1818	0.0551	0.3116	.8702
77-82	1.4865	0.0695	0.4975	.17951	1.3493	0.0849	0.4468	.9025
78-83	1.2453	0.0017	0.0225	.15108	1.088	-0.015	-0.072	.9674
79-84	1.5827	0.084	0.6344	.23290	1.3843	0.0844	0.3979	.9864
80-85	1.8936	0.1499	0.9654	.22750	1.6945	0.1687	0.7593	.9353
81-86	1.6391	0.1257	0.8374	.06589	1.5140	0.1499	0.6236	.8904
82-87	1.9441	0.2024	1.6407	.43078	1.7212	0.2264	0.8685	.8527
83-88	1.6423	0.1185	1.03	.20932	1.5617	0.1283	0.7442	.8175
84-89	1.4199	0.0844	0.6983	.20100	1.3219	0.0849	0.500	.8819
85-90	1.5554	0.0928	0.7444	.23234	1.3790	0.0831	0.4967	.8823
86-91	.93892	-0.0032	-0.1396	.02746	.90174	-0.0027	-0.0168	.9186
87-92	.67602	-0.039	-0.644	-.01351	.7921	-0.021	-0.133	.9254
88-93	.77214	-0.023	-0.448	-.09170	.87509	-0.009	-0.045	.9205

TABLE 2
AVERAGE FUND DIVERSIFICATION SCORE

Period	$S_p \sigma_m / T_p$
75-80	0.565
76-81	0.434
77-82	0.735
78-83	0.363
79-84	0.624
80-85	0.699
81-86	0.624
82-87	0.473
83-88	0.667
84-89	0.711
85-90	0.745
86-91	0.143
87-92	0.384
88-93	0.257

This result is summarised in Table 3, where the frequency of significant values for α_p is insignificantly different from that which might be expected by chance for all periods except 1982-87 and 1983-89. Our interpretation is that the chosen index was itself mean-variance inefficient, thereby

allowing a distribution of funds around the empirical Capital Market and Security Market lines (for the Sharpe and Treynor measures respectively) with very few funds attaining statistically significant risk-adjusted excess return in any given period.

(iii) Fund Ranking Correlations

A comparison of the rankings by the Sharpe and Treynor measures reveals a very high degree of correlation between them, as measured by Spearman's rank-order correlation coefficient [see Kendall, 1970]. These results are given in Table 4. The coefficients are all very high, although there is some variation across periods. Apparently the choice of risk measure is largely unimportant for ranking purposes.

(iv) Persistence of Fund Rankings

An assessment of changes in fund rankings over time is of central importance in fund performance evaluation. Given the high degree of correlation between the two measures we here confine ourselves to a discussion of the Sharpe measure (chosen because the ranking is invariant with respect to the choice of index).

The first column of Table 5 contains the 14 rolling periods of the main analysis. The correlation between the Sharpe rankings for each of these periods and all subsequent non-overlapping periods is given in the body of the table. Cells on the diagonal show the correlation between adjacent non-overlapping five-year periods, while cells to the right show correlations between non-adjacent non-overlapping periods.

TABLE 3
NUMBERS OF FUNDS WITH SIGNIFICANT JENSEN MEASURE

Period	$\alpha > 0$	Significance		$\alpha < 0$	Significance	
		1%	10%		5%	10%
1	27	-	4	65	-	4
2	56	1	-	36	-	-
3	52	-	-	40	-	-
4	58	1	-	34	2	1
5	68	-	3	24	1	-
6	71	1	3	21	-	-
7	63	2	-	29	1	-
8	77	4	8	15	1	1
9	61	5	4	31	1	-
10	59	1	6	33	1	-
11	71	1	3	21	2	-
12	55	1	-	37	3	1
13	40	1	-	52	1	3
14	51	2	4	41	4	2

TABLE 4
CORRELATION BETWEEN THE SHARPE AND TREYNOR RANKINGS

Period	Spearman's ρ
75-80	0.81
76-81	0.96
77-82	0.95
78-83	0.98
79-84	0.96
80-85	0.94
81-86	0.96
82-87	0.86
83-88	0.94
84-89	0.95
85-90	0.94
86-91	0.99
87-92	0.98
88-93	0.99

(All values of ρ are significant at the 0.001% level.)

TABLE 5
INTER-TEMPORAL CONSISTENCY OF SHARPE RANKINGS - ALL TRUSTS

Period	80-85	81-86	82-87	83-88	84-89	85-90	86-91	87-92	88-93
75-79	-.200	-.120	-.069	.191	.137	-.094	-.041	-.045	-.275**
76-80		-.099	-.033	.249*	.191	.016	.029	-.001	-.187
77-81			.157	.392***	.387***	.292**	.211*	.184	.050
78-82				.209*	.242*	.319**	.302**	.392***	.372***
79-83					.169	.270**	.230*	.349***	.434***
80-84						.256*	.205*	.327**	.434***
81-85							.220*	.320**	.467***
82-86								.350***	.250*
83-87									-.026

($n=93$ [92 trusts plus the market index]. 5%, 1% and 0.1% significance denoted by *, ** and *** respectively, using the standard normal distribution - see Kendall, 1970.)

The results are somewhat unexpected. Under the Efficient Markets hypothesis and the assumptions of the CAPM, we should expect no persistence in fund rankings across non-overlapping periods. The reader should recall that the magnitudes of these rank-order correlations depend heavily on sample size, and attention should therefore be paid to their statistical significance. While the diagonal cells of Table 5 apparently show

low correlations between non-overlapping periods, there is a surprising degree of significant correlation from 1977–82 on. That correlation should persist once it occurs is not surprising (successive row or column cells overlap by four out of five years) but we should not expect it to occur in the first place. An explanation for this may be the existence of particular groups of funds which persistently out-perform other groups. We therefore measured inter-temporal correlations within four categories⁷ provided by Datastream – *General*, *Capital Growth*, *Smaller Company* and *Income* Trusts – under the working assumption of homogeneous dividend policy within each group.

Tables 6, 7 and 8 show that the marked degree of significance in the overall intertemporal correlations shown in Table 5 drops sharply for the *General*, *Capital* and *Small Company* categories. However, correlations are generally positive and there remain a small number of statistically significant values which are not easily explained. With respect to the *Income* category, the disaggregation reveals a marked degree of significant intertemporal correlation, shown in Table 9, suggesting that the overall pattern may be largely driven by the persistence of *Income* fund rankings. It seems either that dividend policy was not homogeneous across these funds or that the Datastream classification is insufficiently precise.

TABLE 6
INTER-TEMPORAL CONSISTENCY OF SHARPE RANKINGS – GENERAL TRUSTS

Period	80-85	81-86	82-87	83-88	84-89	85-90	86-91	87-92	88-93
75-79	-.127	.002	.083	.434*	.362*	.122	.073	-.015	-.236
76-80		.144	.077	.379*	.291	.111	.085	.122	-.034
77-81			.012	.443*	.328	.130	.127	.096	-.034
78-82				.088	.190	.250	.328	.371*	.700***
79-83					-.001	.068	.168	.242	.332
80-84						-.146	.0094	.072	.357*
81-85							.078	.116	.340
82-86								.200	.233
83-87									-.109

($n=32$, *, ** and *** denote 5%, 1% and 0.1% significance respectively using the standard normal distribution.)

TABLE 7
INTER-TEMPORAL CONSISTENCY OF SHARPE RANKINGS – CAPITAL TRUSTS

Period	80-85	81-86	82-87	83-88	84-89	85-90	86-91	87-92	88-93
75-79	.006	-.066	-.055	.035	-.155	-.211	-.082	-.255	.038
76-80		-.038	.171	.323	.255	.130	.176	.007	.063
77-81			.185	.213	.284	.228	.256	.232	.145
78-82				.162	.195	.248	.338	.388*	.193
79-83					.178	.252	.242	.486**	.423*
80-84						.258	.216	.581**	.559**
81-85							.092	.315	.492**
82-86								.484*	.141
83-87									.008

($n=25$, * ** and *** denote 5%, 1% and 0.1% significance respectively using the standard normal distribution.)

TABLE 8
INTER-TEMPORAL CONSISTENCY OF SHARPE RANKINGS – SMALL COMPANY TRUSTS

Period	80-85	81-86	82-87	83-88	84-89	85-90	86-91	87-92	88-93
75-79	-.188	-.042	.370	.406	.503	.333	.382	.867*	.430
76-80		-.115	.273	.455	.479	.273	.309	.745	.333
77-81			.345	.406	.491	.297	.455	.818*	.455
78-82				-.176	-.067	.030	.321	.673	.552
79-83					-.030	.091	.345	.297	.321
80-84						.006	.212	.079	.345
81-85							.491	.079	.588
82-86								.539	.552
83-87									.091

(For $n=8$, * indicates 5% significance – see exact probabilities for ρ in Kendall, 1970.)

TABLE 9
INTER-TEMPORAL CONSISTENCY OF SHARPE RANKINGS – INCOME TRUSTS

Period	80-85	81-86	82-87	83-88	84-89	85-90	86-91	87-92	88-93
75-79	-.227	-.055	-.014	.055	.157	-.001	.147	.116	-.411*
76-80		-.319	-.193	-.140	-.090	-.182	-.049	-.269	-.499*
77-81			.309	.403*	.558**	.659***	.496*	.305	.445*
78-82				.407*	.551**	.664***	.379	.464*	.566**
79-83					.524**	.552**	.274	.447*	.714***
80-84						.534**	.215	.458*	.606***
81-85							.288	.535**	.488*
82-86								.435*	.426*
83-87									.347

($n=25$, * ** and *** denote 5%, 1% and 0.1% significance respectively using the standard normal distribution.)

We offer a number of possible reasons for the persistence of fund rankings.

First, it may very well be that particular managers do have superior ability in picking high-performing stocks – a possibility not permitted by the CAPM or the efficient markets hypothesis. We note that certain funds actually showed very persistent relative performance. Table 10 gives examples from the *Income* category, where for the three successive non-overlapping periods 1978–82, 1983–87, and 1988–93 we see that 10 out of 25 funds retained their relative ranking to within five rank places. Some funds also showed persistent positive or negative performance. In Table 11 we have counted the numbers of funds with 11 Jensen scores of the same sign out of a possible 14, identifying those with at least three significant Jensen scores. Here we see that 29 funds scored persistently positively (good funds) and 7 funds persistently badly (bad funds) while three each of the good and bad fund groups managed three or more significant Jensen scores. The bad funds with one exception are from the income group, suggesting again that correction for dividend payments may matter in comparisons of this group with other groups.

Second, the Mean-Variance/CAPM framework may provide an inappropriate framework for risk adjustment – the fund ranking could very well persist if funds were earning returns attributable to undetected sources of systematic risk. A multi-index model or the Arbitrage Pricing model of Ross [1978] might provide a more sensitive benchmark [see Grinblatt and Titman, 1992, for example].

Finally, we speculate that there may be cyclical economic conditions which favour the activities of particular fund managers at different times, and which therefore recreate particular patterns of fund performance.

TABLE 10
INCOME FUNDS WITH RELATIVE RANK UNCHANGED BEYOND FIVE RANK PLACES IN
SUCCESSIVE NON-OVERLAPPING PERIODS

Fund Number	Fund Rank out of 25		
	75-79	80-84	85-89
17	13	25	23
21	8	20	24
47	24	21	26
57	3	1	5
59	9	23	19
61	11	19	22
63	17	10	8
67	23	3	7
68	21	17	12
69	18	4	9
78	20	22	21
83	22	8	4
88	5	14	10
90	15	11	15

Fund Number	Fund Rank out of 25		
	78-82	83-87	88-92
7	21	18	16
11	6	15	11
17	22	25	25
23	19	24	21
27	9	11	9
46	10	9	10
59	20	23	20
61	17	26	22
63	14	8	6
66	7	5	5
67	4	4	7
78	23	21	18
84	24	19	19
89	15	6	3

TABLE 11.
NUMBERS OF FUNDS PERFORMING WELL OR BADLY, BY FUND CATEGORY

Number of positively-performing funds				
<i>General</i>	<i>Capital</i>	<i>Income</i>	<i>Small</i>	<i>Total</i>
12 (1)	11 (2)	5	1	29

Number of negatively-performing funds				
<i>General</i>	<i>Capital</i>	<i>Income</i>	<i>Small</i>	<i>Total</i>
0	1	7 (3)	0	8

(Funds cited in Table 11 have 11 out of the possible 14 Jensen scores with the same sign. The number of funds with 3 or more significant scores out of 14 is given in brackets.)

CONCLUSION

With regard to the average performance of funds, our results suggest that investment trusts have not on average been able to out-perform the market index when adjusted for risk using single-index adjustment models. To the extent that there is a survivorship bias, due to the deletion of de-listed funds from our Datastream database, this poor performance would be further weakened. We have also shown that the choice of variance or covariance risk (respectively the Sharpe and Treynor measures) seems to matter very little.

The performance of individual funds is more variable. The various significant positive inter-temporal rank-order correlations seem to suggest that some fund managers may be able to beat the market quite consistently on a risk-adjusted basis. Whether or not this is evidence of market inefficiency (trading on superior information) is debatable. The result is also consistent with the view that some managers may be compensated for obtaining and processing costly information. If stock markets were fully efficient and prices reflected all available information then fund managers would not be able to earn excess returns by spending resources on research. However, when information is costly the existence of a small number of informed traders can generate a wedge between trade prices and full-information prices which is sufficient to compensate professional arbitrage trading [see Grossman, 1976; Ippolito, 1989]. Through the compensated actions of arbitrageurs the market reaches full-information equilibrium prices. In such a model passive investors essentially pay for the

information-gathering activity, while informed traders, who may appear to 'beat the market' before expenses, earn no excess returns after the expense of gathering information is netted out. It would be necessary to identify and adjust for fund-management costs in order to eliminate this possibility – a task well beyond the scope of this article.

NOTES

1. It is not usual to find relative ranking claims which are risk-adjusted.
2. Prices are from Datastream International. Dividend-adjustment is not possible using this dataset but since this is only relevant for observation dates which coincide with ex-dividend dates, we believe the bias to be small. Comparisons between performance measures for the same funds are in any case legitimate.
3. No data were available for funds which ceased trading before 1993, implying survivorship bias. Correction would further weaken the poor overall fund performance.
4. Disaggregated results for all 92 trusts over all periods are available on request.
5. This restriction, $E(\epsilon_i \epsilon_j) = 0, \forall i \neq j$, asserts that the market model residuals are uncorrelated across assets.
6. The statistical significance of the difference between $S_p \sigma_m / T_p$ and unity is not readily assessable, since the distribution is unknown, but increasing departures from unity constitute decreasing diversification.
7. The investments trusts may fall into more than one category. We ignore international categories, since funds in these groups fall into the four categories identified.

REFERENCES

- Admati, R.A. and S.A. Ross, 1985, 'Measuring Investment Performance in a Rational Expectations Equilibrium Model', *Journal of Business*, Vol.58, pp.1-26.
- Eun, C., R. Kolodny and B.G. Resnik, 1991, 'US-Based International Mutual Funds: A Performance Evaluation', *The Journal of Portfolio Management*, pp.88-94.
- Cheney, J.M., S. Atkinson and B.A. Bailey, 1992, 'International Mutual Fund Performance US vs. UK', *Managerial Finance*, 18, pp.39-48.
- Coggin, T.D., F.J. Fabozzi and S. Rahman, 1993, 'The Investment Performance of US Equity Pension Fund Managers: An Empirical Investigation', *Journal of Finance*, XLVIII, pp.1038-55.
- Connor, G. and R.A. Korajczyk, 1986, 'Performance Measurement with the Arbitrage Pricing Theory: A New Framework for Analysis', *Journal of Financial Economics*, 15, pp.374-94.
- Cumby, R.E. and J.D. Glen, 1990, 'Evaluating the Performance of International Mutual Funds', *The Journal of Finance*, XLV, pp.497-521.
- Dybvig, P.H. and S.A. Ross, 1985, 'The Analytics of Performance Measurement Using a Security Market Line', *Journal of Finance*, XL, pp.401-15.
- Grant, D., 1977, 'Portfolio Performance and the Cost of Timing Decisions', *Journal of Finance*, 32, pp.837-46.
- Grimblatt M. and S. Titman, 1989, 'Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings', *Journal of Business*, 62, pp.393-416.
- Grimblatt M. and S. Titman, 1992, 'The Persistence of Mutual Fund Performance', *Journal of Finance*, XLVII, 1977-84.
- Henriksson, R.D., 1984, Market Timing and Mutual Fund Performance: An Empirical Investigation', *Journal of Business*, 57, pp.73-96.

- Ippolito R.A., 1989, 'Efficiency with Costly Information: A Study of Mutual Fund Performance, 1965-1984', *Quarterly Journal of Economics*, CIV, pp.1-23.
- Jensen M.C., 1968, 'The Performance of Mutual Funds in the Period 1945-1964', *Journal of Finance*, 23, pp.389-416.
- Kendall, M.G., 1970, *Rank Correlation Methods*, fourth edition, London: Griffin.
- Kon, S.J., 1983, 'The Market Timing Performance of Mutual Fund Managers', *Journal of Business*, 56, pp.323-47.
- Kon, S.J. and F.C. Jen, 1979, 'The Investment Performance of Mutual Funds: An Empirical Investigation of Timing, Selectivity and Market Efficiency', *Journal of Business*, 52, pp.263-89.
- Lakonishok, J., A. Shleifer and R.W. Vishny, 1992, 'The Structure and Performance of the Money Management Industry', *Brookings Papers on Economic Activity: Microeconomics*, pp.339-79.
- Lehmann, B.N. and D.M. Modest, 1987, 'Mutual Fund Performance Evaluation: A Comparison of Benchmarks and Benchmark Comparisons', *Journal of Finance*, 42, pp.233-65.
- Roll, R., 1977, 'A Critique of the Asset Pricing Theory's Tests; Part I: On Past and Potential Testability of the Theory', *Journal of Financial Economics*, 4, pp.129-76.
- Roll, R., 1978, 'Ambiguity When Performance is Measured by the Security Market Line', *Journal of Finance*, 33, pp.1051-69.
- Ross, A.S., 1978, 'The Arbitrage Theory of Capital Asset Pricing', *Journal of Economic Theory*, 13, pp.341-60.
- Sharpe, W., 1966, 'Mutual Fund Performance', *Journal of Business*, 34, pp.119-38.
- Stambaugh, R.F., 1982, 'On the Exclusion of Assets from Tests of the Two-Parameter Model: A Sensitivity Analysis', *Journal of Financial Economics*, 10, pp.237-68.
- Treynor, J.L., 1965, 'How to Rate Management of Investment Funds', *Harvard Business Review*, 43, pp.63-75.
- Treynor, J.L. and K.K. Mazuy, 1967, 'Can Mutual Funds Outguess the Market?' *Harvard Business Review*, 44, pp.131-6.

